RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _09/830,502C_

Source: _IPW16_

Date Processed by STIC: _1-13-05_

ENTERED

RAW SEQUENCE LISTING   DATE: 01/13/2005
PATENT APPLICATION: US/09/830,502C   TIME: 17:23:47

Input Set : A:\C26151.app
Output Set: N:\CRF4\01132005\I830502C.raw

```
 3 <110> APPLICANT: Barany, Francis
 4        Cao, Weiguo
 5        Tong, Jie
 7 <120> TITLE OF INVENTION: HIGH FIDELITY THERMOSTABLE LIGASE AND USES THEREOF
 9 <130> FILE REFERENCE: 19603/2615
11 <140> CURRENT APPLICATION NUMBER: 09/830,502C
12 <141> CURRENT FILING DATE: 2001-10-09
14 <150> PRIOR APPLICATION NUMBER: 60/106,461
15 <151> PRIOR FILING DATE: 1998-10-30
17 <150> PRIOR APPLICATION NUMBER: PCT/US99/25437
18 <151> PRIOR FILING DATE: 1999-10-29
20 <160> NUMBER OF SEQ ID NOS: 24
22 <170> SOFTWARE: PatentIn Ver. 2.1
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 674
26 <212> TYPE: PRT
27 <213> ORGANISM: Thermus sp.
29 <400> SEQUENCE: 1
30 Met Thr Leu Glu Glu Ala Arg Arg Arg Val Asn Glu Leu Arg Asp Leu
31  1               5                  10                  15
33 Ile Arg Tyr His Asn Tyr Leu Tyr Tyr Val Leu Asp Ala Pro Glu Ile
34              20                  25                  30
36 Ser Asp Ala Glu Tyr Asp Arg Leu Leu Arg Glu Leu Lys Glu Leu Glu
37          35                  40                  45
39 Glu Arg Phe Pro Glu Leu Lys Ser Pro Asp Ser Pro Thr Glu Gln Val
40      50                  55                  60
42 Gly Ala Arg Pro Leu Glu Ala Thr Phe Arg Pro Val Arg His Pro Thr
43  65              70                  75                  80
45 Arg Met Tyr Ser Leu Asp Asn Ala Phe Ser Leu Asp Glu Val Arg Ala
46              85                  90                  95
48 Phe Glu Glu Arg Ile Glu Arg Ala Leu Gly Arg Lys Gly Pro Phe Leu
49             100                 105                 110
51 Tyr Thr Val Glu Arg Lys Val Asp Gly Leu Ser Val Asn Leu Tyr Tyr
52          115                 120                 125
54 Glu Glu Gly Ile Leu Val Phe Gly Ala Thr Arg Gly Asp Gly Glu Thr
55      130                 135                 140
57 Gly Glu Glu Val Thr Gln Asn Leu Leu Thr Ile Pro Thr Ile Pro Arg
58 145                 150                 155                 160
60 Arg Leu Thr Gly Val Pro Asp Arg Leu Glu Val Arg Gly Glu Val Tyr
61              165                 170                 175
63 Met Pro Ile Glu Ala Phe Leu Arg Leu Asn Gln Glu Leu Glu Glu Ala
64             180                 185                 190
66 Gly Glu Arg Ile Phe Lys Asn Pro Arg Asn Ala Ala Ala Gly Ser Leu
```

```
67          195                200                205
69 Arg Gln Lys Asp Pro Arg Val Thr Ala Arg Arg Gly Leu Arg Ala Thr
70     210                215                220
72 Phe Tyr Ala Leu Gly Leu Gly Leu Glu Glu Thr Gly Leu Lys Ser Gln
73 225                230                235                240
75 His Asp Leu Leu Leu Trp Leu Arg Glu Arg Gly Phe Pro Val Glu His
76                245                250                255
78 Gly Phe Thr Arg Ala Leu Gly Ala Glu Gly Val Glu Glu Val Tyr Gln
79                260                265                270
81 Ala Trp Leu Lys Glu Arg Arg Lys Leu Pro Phe Glu Ala Asp Gly Val
82     275                280                285
84 Val Val Lys Leu Asp Asp Leu Ala Leu Trp Arg Glu Leu Gly Tyr Thr
85     290                295                300
87 Ala Arg Thr Pro Arg Phe Ala Leu Ala Tyr Lys Phe Pro Ala Glu Glu
88 305                310                315                320
90 Lys Glu Thr Arg Leu Leu Ser Val Ala Phe Gln Val Gly Arg Thr Gly
91                325                330                335
93 Arg Ile Thr Pro Val Gly Val Leu Glu Pro Val Phe Ile Glu Gly Ser
94                340                345                350
96 Glu Val Ser Arg Val Thr Leu His Asn Glu Ser Phe Ile Glu Glu Leu
97                355                360                365
99 Asp Val Arg Ile Gly Asp Trp Val Leu Val His Lys Ala Gly Gly Val
100     370                375                380
102 Ile Pro Glu Val Leu Arg Val Leu Lys Glu Arg Arg Thr Gly Glu Glu
103 385                390                395                400
105 Lys Pro Ile Ile Trp Pro Glu Asn Cys Pro Glu Cys Gly His Ala Leu
106                405                410                415
108 Ile Lys Glu Gly Lys Val His Arg Cys Pro Asn Pro Leu Cys Pro Ala
109                420                425                430
111 Lys Arg Phe Glu Ala Ile Arg His Tyr Ala Ser Arg Lys Ala Met Asp
112     435                440                445
114 Ile Gln Gly Leu Gly Glu Lys Leu Ile Glu Lys Leu Leu Glu Lys Gly
115     450                455                460
117 Leu Val Arg Asp Val Ala Asp Leu Tyr Arg Leu Lys Lys Glu Asp Leu
118 465                470                475                480
120 Val Asn Leu Glu Arg Met Gly Glu Lys Ser Ala Glu Asn Leu Leu Arg
121                485                490                495
123 Gln Ile Glu Glu Ser Lys Gly Arg Gly Leu Glu Arg Leu Leu Tyr Ala
124                500                505                510
126 Leu Gly Leu Pro Gly Val Gly Glu Val Leu Ala Arg Asn Leu Ala Leu
127     515                520                525
129 Arg Phe Gly His Met Asp Arg Leu Leu Glu Ala Gly Leu Glu Asp Leu
130     530                535                540
132 Leu Glu Val Glu Gly Val Gly Glu Leu Thr Ala Arg Ala Ile Leu Asn
133 545                550                555                560
135 Thr Leu Lys Asp Pro Glu Phe Arg Asp Leu Val Arg Arg Leu Lys Glu
136                565                570                575
138 Ala Gly Val Glu Met Glu Ala Lys Glu Arg Glu Gly Glu Ala Leu Lys
139     580                585                590
```

```
141 Gly Leu Thr Phe Val Ile Thr Gly Glu Leu Ser Arg Pro Arg Glu Glu
142         595                 600                 605
144 Val Lys Ala Leu Leu Arg Arg Leu Gly Ala Lys Val Thr Asp Ser Val
145         610                 615                 620
147 Ser Arg Lys Thr Ser Phe Leu Val Val Gly Glu Asn Pro Gly Ser Lys
148 625                 630                 635                 640
150 Leu Glu Lys Ala Arg Ala Leu Gly Val Pro Thr Leu Ser Glu Glu Glu
151             645                 650                 655
153 Leu Tyr Arg Leu Ile Glu Glu Arg Thr Gly Lys Asp Pro Arg Ala Leu
154         660                 665                 670
156 Thr Ala
160 <210> SEQ ID NO: 2
161 <211> LENGTH: 2025
162 <212> TYPE: DNA
163 <213> ORGANISM: Thermus sp.
165 <400> SEQUENCE: 2
166 atgaccctag aggaggcccg caggcgcgtc aacgaactca gggacctgat ccgttaccac 60
167 aactacctct attacgtctt ggacgccccc gagatctccg acgccgagta cgaccggctc 120
168 cttagggagc ttaaggagct ggaggagcgc tttcccgagc tcaaaagccc cgactccccc 180
169 acggaacagg tggggcgag gcctctggag gccaccttcc gcccggtgcg ccaccccacc 240
170 cgcatgtact ccctggacaa cgccttttcc ttggacgagg tgagggcctt tgaggagcgc 300
171 atagagcggg ccctggggcg gaaggggccc ttcctctaca ccgtggagcg caaggtggac 360
172 ggtctttccg tgaacctcta ctacgaggag ggcatcctcg tctttggggc caccccggggc 420
173 gacggggaga ccggggagga ggtgacccag aacctcctca ccatccccac cattccccgc 480
174 cgcctcacgg gcgttccgga ccgcctcgag gtccggggcg aggtctacat gcccatagag 540
175 gccttcctca ggctcaacca ggagctggag gaggcggggg agcgcatctt caaaaacccc 600
176 aggaacgccg ccgccgggtc cttgcggcag aaagacccca gggtcacggc caggcggggc 660
177 ctgagggcca ccttttacgc cctggggctg gccctggagg aaaccgggtt aaaaagccag 720
178 cacgacctтc tcctatggct aagagagcgg ggctttcccg tggagcacgg ctttacccgg 780
179 gccctggggg cggaggggt ggaggaggtc taccaggcct ggctcaagga gaggcggaag 840
180 cttccctttg aggccgacgg ggtggtggtc aagctggacg acctcgccct ctggcgggag 900
181 ctggggtaca ccgcccgcac ccccgcttc gccctcgcct acaagttccc ggccgaggag 960
182 aaggagaccc gcctcctctc cgtggccttc caggtggggc ggaccgggcg catcacccc 1020
183 gtgggcgttc tggagcccgt cttcatagag ggcagcgagg tgagccgggt cacccтccac 1080
184 aacgagagct tcattgagga gctggacgtg cgcatcggcg actgggtgct ggtccacaag 1140
185 gcgggcgggg tgattcccga ggtgctgagg gtcctgaaag agcgccgcac cggggaggag 1200
186 aagcccatca tctggcccga gaactgcccc gagtgcggcc acgccctcat caaggagggg 1260
187 aaggtccacc gctgccccaa ccccttgtgc cccgccaagc gctttgaggc catccgccac 1320
188 tacgcctccc gcaaggccat ggacatccag ggcctggggg agaagctcat agaaaagctt 1380
189 ctggaaaagg gcctggtccg ggacgtggcc gacctctacc gcctgaagaa ggaggacctg 1440
190 gtgaacctgg agcgcatggg ggagaagagc gcagagaacc tcctccgcca gatagaggag 1500
191 agcaagggcc gcggcctgga gcgcctcctt tacgccctgg gccttcccgg ggtggggggag 1560
192 gtgctggccc ggaacctggc cctccgcttc ggccacatgg accgccttct ggaggcgggc 1620
193 ctcgaggacc tcctggaggt ggaggggggtg ggcgagctca ccgccgggc catcctgaat 1680
194 accctaaagg acccggagtt ccgggacctg gtgcgccgcc tgaaggaggc cggggtggag 1740
195 atggaggcca aagagcggga gggcgaggcc ttgaaggggc tcaccttcgt catcaccggg 1800
196 gagctttccc ggccccggga ggaggtgaag gccctcctta ggcggcttgg ggccaaggtg 1860
197 acggactcgg tgagccgcaa gacgagcttc ctggtggtgg gggagaaccc ggggagcaag 1920
198 ctggaaaagg cccgcgcctt ggggggtcccc accctgagcg aggaggagct ctaccgcctc 1980
```

```
199 attgaggaga ggacgggcaa ggacccaagg gccctcacgg cctag                    2025
202 <210> SEQ ID NO: 3
203 <211> LENGTH: 20
204 <212> TYPE: DNA
205 <213> ORGANISM: Artificial Sequence
207 <220> FEATURE:
208 <223> OTHER INFORMATION: Description of Artificial Sequence: probe or
209       primer
211 <220> FEATURE:
212 <221> NAME/KEY: tRNA
213 <222> LOCATION: (4)
214 <223> OTHER INFORMATION: w at position 4 can be T or A
216 <220> FEATURE:
217 <221> NAME/KEY: unsure
218 <222> LOCATION: (5)
219 <223> OTHER INFORMATION: s at position 5 can be C or G
221 <220> FEATURE:
222 <221> NAME/KEY: unsure
223 <222> LOCATION: (12)
224 <223> OTHER INFORMATION: s at position 12 can be C or G
226 <220> FEATURE:
227 <221> NAME/KEY: unsure
228 <222> LOCATION: (15)
229 <223> OTHER INFORMATION: r at position 15 can be G or A
231 <220> FEATURE:
232 <221> NAME/KEY: unsure
233 <222> LOCATION: (18)
234 <223> OTHER INFORMATION: y at position 18 can be T or C
236 <400> SEQUENCE: 3
237 atcwscgacg csgartayga                                                 20
240 <210> SEQ ID NO: 4
241 <211> LENGTH: 7
242 <212> TYPE: PRT
243 <213> ORGANISM: Artificial Sequence
245 <220> FEATURE:
246 <223> OTHER INFORMATION: Description of Artificial Sequence: protein
248 <400> SEQUENCE: 4
249 Ile Ser Asp Ala Glu Tyr Asp
250   1                 5
253 <210> SEQ ID NO: 5
254 <211> LENGTH: 20
255 <212> TYPE: DNA
256 <213> ORGANISM: Artificial Sequence
258 <220> FEATURE:
259 <223> OTHER INFORMATION: Description of Artificial Sequence: probe or
260       primer
262 <220> FEATURE:
263 <221> NAME/KEY: unsure
264 <222> LOCATION: (3)
```

RAW SEQUENCE LISTING   DATE: 01/13/2005
PATENT APPLICATION: **US/09/830,502C** TIME: 17:23:47

Input Set : **A:\C26151.app**
Output Set: **N:\CRF4\01132005\I830502C.raw**

```
265 <223> OTHER INFORMATION: s at position 3 can be C or G
267 <220> FEATURE:
268 <221> NAME/KEY: unsure
269 <222> LOCATION: (6)
270 <223> OTHER INFORMATION: s at position 6 can be C or G
272 <220> FEATURE:
273 <221> NAME/KEY: unsure
274 <222> LOCATION: (8)
275 <223> OTHER INFORMATION: k at position 8 can be G or T
277 <220> FEATURE:
278 <221> NAME/KEY: unsure
279 <222> LOCATION: (9)
280 <223> OTHER INFORMATION: s at position 9 can be G or C
282 <220> FEATURE:
283 <221> NAME/KEY: unsure
284 <222> LOCATION: (12)
285 <223> OTHER INFORMATION: s at position 12 can be G or C
287 <220> FEATURE:
288 <221> NAME/KEY: unsure
289 <222> LOCATION: (15)
290 <223> OTHER INFORMATION: y at position 15 can be C or T
292 <220> FEATURE:
293 <221> NAME/KEY: unsure
294 <222> LOCATION: (18)
295 <223> OTHER INFORMATION: r at position 18 can be A or G
297 <400> SEQUENCE: 5
298 ccsgtscksc csacytgraa                                    20
301 <210> SEQ ID NO: 6
302 <211> LENGTH: 20
303 <212> TYPE: DNA
304 <213> ORGANISM: Artificial Sequence
306 <220> FEATURE:
307 <223> OTHER INFORMATION: Description of Artificial Sequence: probe or
308       primer
310 <220> FEATURE:
311 <221> NAME/KEY: unsure
312 <222> LOCATION: (9)
313 <223> OTHER INFORMATION: v at position 9 can be C, G, or A
315 <220> FEATURE:
316 <221> NAME/KEY: unsure
317 <222> LOCATION: (11)
318 <223> OTHER INFORMATION: r at position 11 can be A or G
320 <220> FEATURE:
321 <221> NAME/KEY: unsure
322 <222> LOCATION: (12)
323 <223> OTHER INFORMATION: y at position 12 can be T or C
325 <220> FEATURE:
326 <221> NAME/KEY: unsure
327 <222> LOCATION: (16)
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 01/13/2005
PATENT APPLICATION: US/09/830,502C   TIME: 17:23:48

Input Set : **A:\C26151.app**
Output Set: **N:\CRF4\01132005\I830502C.raw**

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

```
Seq#:15; Xaa Pos. 18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36
Seq#:15; Xaa Pos. 37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55
Seq#:15; Xaa Pos. 56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74
Seq#:15; Xaa Pos. 75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93
Seq#:15; Xaa Pos. 94,95,96,97,98,99,100,101,102,103,104,105,106,107,108,109
Seq#:15; Xaa Pos. 110,111,112,113,114,115,116,117,118,119,120,126,127,128
Seq#:15; Xaa Pos. 129,130,131,132,133,134,135,136,137,138,139,140,141,142
Seq#:15; Xaa Pos. 143,144,145,146,147,148,149,150,151,152,153,154,155,156
Seq#:15; Xaa Pos. 157,158,159,160,161,162,163,164,165,166,167,168,169,170
Seq#:15; Xaa Pos. 171,172
Seq#:16; Xaa Pos. 18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36
Seq#:16; Xaa Pos. 37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55
Seq#:16; Xaa Pos. 56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74
Seq#:16; Xaa Pos. 75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93
Seq#:16; Xaa Pos. 94,95,96,97,98,99,100,101,102,103,104,105,106,107,108,109
Seq#:16; Xaa Pos. 110,111,112,113,114,115,116,117,118,119,120,129,130,131
Seq#:16; Xaa Pos. 132,133,134,135,136,137,138,139,140,141,142,143,144,145
Seq#:16; Xaa Pos. 146,147,148,149,150,151,152,153,154,155,156,157,158,159
Seq#:16; Xaa Pos. 160,161,162,163,164,165,166,167,168,169,170,171,172,173
Seq#:16; Xaa Pos. 174,175
Seq#:17; Xaa Pos. 18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36
Seq#:17; Xaa Pos. 37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55
Seq#:17; Xaa Pos. 56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74
Seq#:17; Xaa Pos. 75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93
Seq#:17; Xaa Pos. 94,95,96,97,98,99,100,101,102,103,104,105,106,107,108,109
Seq#:17; Xaa Pos. 110,111,112,113,114,115,116,117,118,119,120,126,127,128
Seq#:17; Xaa Pos. 129,130,131,132,133,134,135,136,137,138,139,140,141,142
Seq#:17; Xaa Pos. 143,144,145,146,147,148,149,150,151,152,153,154,155,156
Seq#:17; Xaa Pos. 157,158,159,160,161,162,163,164,165,166,167,168,169,170
Seq#:17; Xaa Pos. 171,172
Seq#:18; Xaa Pos. 18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36
Seq#:18; Xaa Pos. 37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55
Seq#:18; Xaa Pos. 56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74
Seq#:18; Xaa Pos. 75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93
Seq#:18; Xaa Pos. 94,95,96,97,98,99,100,101,102,103,104,105,106,107,108,109
Seq#:18; Xaa Pos. 110,111,112,113,114,115,116,117,118,119,120,126,127,128
Seq#:18; Xaa Pos. 129,130,131,132,133,134,135,136,137,138,139,140,141,142
Seq#:18; Xaa Pos. 143,144,145,146,147,148,149,150,151,152,153,154,155,156
Seq#:18; Xaa Pos. 157,158,159,160,161,162,163,164,165,166,167,168,169,170
Seq#:18; Xaa Pos. 171,172
Seq#:19; Xaa Pos. 18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36
Seq#:19; Xaa Pos. 37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55
Seq#:19; Xaa Pos. 56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74
Seq#:19; Xaa Pos. 75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 01/13/2005
PATENT APPLICATION: **US/09/830,502C**   TIME: 17:23:48

Input Set : **A:\C26151.app**
Output Set: **N:\CRF4\01132005\I830502C.raw**

```
Seq#:19; Xaa Pos. 94,95,96,97,98,99,100,101,102,103,104,105,106,107,108,109
Seq#:19; Xaa Pos. 110,111,112,113,114,115,116,117,118,119,120,126,127,128
Seq#:19; Xaa Pos. 129,130,131,132,133,134,135,136,137,138,139,140,141,142
Seq#:19; Xaa Pos. 143,144,145,146,147,148,149,150,151,152,153,154,155,156
Seq#:19; Xaa Pos. 157,158,159,160,161,162,163,164,165,166,167,168,169,170
Seq#:19; Xaa Pos. 171,172
Seq#:20; Xaa Pos. 18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36
```

VERIFICATION SUMMARY   DATE: 01/13/2005
PATENT APPLICATION: **US/09/830,502C**   TIME: 17:23:48

Input Set : **A:\C26151.app**
Output Set: **N:\CRF4\01132005\I830502C.raw**

```
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:469 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:16
M:341 Repeated in SeqNo=15
L:522 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:16
M:341 Repeated in SeqNo=16
L:575 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17 after pos.:16
M:341 Repeated in SeqNo=17
L:628 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18 after pos.:16
M:341 Repeated in SeqNo=18
L:681 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19 after pos.:16
M:341 Repeated in SeqNo=19
L:734 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20 after pos.:16
M:341 Repeated in SeqNo=20
L:787 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21 after pos.:16
M:341 Repeated in SeqNo=21
L:840 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22 after pos.:16
M:341 Repeated in SeqNo=22
L:893 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23 after pos.:16
M:341 Repeated in SeqNo=23
L:941 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:24 after pos.:0
```